Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2018 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

6-26-2018

Understanding Streaming Music Diffusion in a Semi-Closed Social Environment

Jing Ren
Singapore Management University, jing.ren.2012@phdis.smu.edu.sg

Robert Kauffman
Singapore Management University, rkauffman@smu.edu.sg

Understanding Streaming Music Diffusion in a Semi-Closed Social Environment

Completed Research Paper

Jing Ren
Singapore Management University
80 Stamford Road, Singapore
Jing.ren.2012@phdis.smu.edu.sg

Robert J. Kauffman
Singapore Management University
80 Stamford Road, Singapore
rkauffman@smu.edu.sg

Abstract

Music social networks play a role in the diffusion of music. There are different ways a piece of music reaches people in a network: through the influence of social connections or via the discovery of external information, such as mass media, newspapers, etc. This empirical study uses over 10 months of user listening data from a music social network to examine the effects of external information on streaming music diffusion at the macro- and micro-levels. The data include weekly listening records for 557,554 users. Our results suggest that external information is a significant driver of increased streaming music diffusion, in comparison to in-network influences. We also found evidence of variation in the different influences, such as for a scale effect, the validity and type of information shared, and the impact of geolocation. These insights can be used to promote music and design personalized music recommendations.

Keywords: Causal inference, count data model, difference-in-differences (DiD) model, external information, information discovery, music diffusion, music social networks, social influence.

Introduction

Streaming music has taken the music market by storm, and has diminished the share of physical album sales and digital downloads. The 2017 Global Music Report (IFPI 2017) suggests that, by the end of 2016, physical album sales sunk to an historical low, while gross annual music revenues increased, turning around a decreasing trend since 2000. Although people are less likely to buy CDs now, they still are listening more than ever to digital music. In the U.S., the sales volume of digital music is now around 10 times that of physical music. Compared to digital download consumers, streaming music consumers do not need to download files when they subscribe: instead, they can enjoy music for free when they have access to an Internet connection.[1]

Music streaming via music social networks involves a special technology platform for making the depth and richness of music available to millions of people. The platform supports social connections by introducing artists and music products directly to Internet audiences. This has brought opportunities to the music industry to figure out how social networks diffuse music, and how platforms can strengthen the connection between listeners and artists, thereby effectively promote music in real time (IFPI 2012, 2015). Such insights are critical for effective marketing plans, and to maximize music and artist value. Also, understanding music diffusion can improve social network services to retain loyal users.

This is not just an important business problem for the music industry though. It is also a challenging research question from an academic perspective. It is challenging because streaming diffusion occurs within platforms that represent *semi-closed environments* (Garg et al. 2011). People's listening behavior for a song or an artist may be influenced by other users on the platform through their social connections. It also may be impacted by artist-related information from other out-of-network factors that we refer to as *external information*. These factors include music content news, and artists' social activities on other social networks and mass media platforms, such as radio and TV. The related categories are:

- **Social influence.** This is the effect of social relation or in-network recommendations related to a user's listening behavior, which involves a person's acceptance of what the system can supply.
- **Information discovery**. This is the effect of mass media information or out-of-network news related to a user's listening behavior, for a person to learn about what to buy or consume.

Although social influences impact artists' music diffusion (Bapna and Umyarov 2015, Pálovics and Benczúr 2015), without the external force associated with content promotion, the diffusion usually will be slower and decline over time to a relatively stable level, even for superstars like Adele. This suggests firms may leverage external information for consumer engagement via music social networks.[2]

This research explores the information discovery effects of music diffusion. We contribute new knowledge by focusing on external information that drives diffusion. (1) What kinds of external information affect streaming music diffusion? (2) How large and persistent are the effects of external information? (3) What kinds of information can be used to improve personalized music recommendations?

We applied a *causal inference approach* to assess the effects of external information on streaming music diffusion at the macro- and micro-levels, using a large panel dataset on listening in a music social network. The dataset contains over 557,000 users' weekly listening records, and we tracked user listening for 1,300 artists over about one year of time. To further analyze music diffusion, we collected data on the two categories of external information, as well as artist and user characteristics, and listening behavior (listening taste, demographics, etc.). Our analysis of the effects of external information leverages the use of dependent variables that represent artist music adoption and user listening behavior.

---

[1] Revenues for digital downloads in 2016 dropped to 60% of their 2012 peak in the U.S., while streaming revenues in 2016 doubled from 2015, exceeding digital downloads (RIAA 2016). Expansion of streaming services is continuing, with revenues in 1Q-2Q 2017 accounting for 62% of the total market, with digital downloads achieving a declining 19% share (RIAA 2017).
[2] An example is the burst of interest and listening to American vocalist, Ariana Grande, after the terror attack on her concert in Manchester, Great Britain. Although it was not a suitable time to promote Ariana's music, social listening still reacted positively to external news of this negative event.

Our study finds that external information has a significant impact on an artist's music diffusion in a music social network, and the impact and persistence are related to the details of information that becomes available about the artist. This information discovery effect occurs at both the artist and user levels. We also find that a listener's geolocation limits diffusion. So although people can access whatever music they like, more limited access to external information may constrain their discovery. This research contributes to the literature on diffusion analysis in music social networks, by considering external information within a semi-closed music streaming environment. Prior research has pointed out that there is media influence (Garg et al. 2011, Myers et al. 2012), however, what kind of information can be leveraged to promote artists and their music in social networks remains unclear.

Theory

Research in Computer Science (CS) and Information Systems (IS) has estimated the influence effects on information diffusion in social networks. Some of the issues that have been studied include: the effects of social relations and capital (Ellison et al. 2011, Bapna and Umyarov 2015, Sharma and Cosley 2016); the role of weak and strong ties (Bakshy et al. 2012); the impacts of social recommendations (Garg et al. 2011); and what happens when user-generated content is considered (Susarla et al. 2012). For music diffusion, however, surprisingly conclusions related to how the main influences work have been obtained. For example, Bapna and Umyarov (2015) confirmed the significance of social influence through friends related to music subscriptions. Sharma and Cosley (2016) indicated that the effect of social influence has been over-estimated. They also found that the majority of shared music listening between friends is due to homophily, and not influence. In their findings, less than 1% of users' actions can be explained by their friends' influence. Garg et al. (2011) and Dewan et al. (2017) also reported evidence for music diffusion and social influence, but social influence was more important for music with narrow or niche appeal, compared to more broadly appealing music.

Studies on the information discovery effects on music diffusion have been rare. They are challenging to implement due to the unobservable process involving information discovery related to music search through to consumption (Garg et al. 2011). Music networks operate in semi-closed platform environments, so listeners can discover information from the media and via content sampling. Internet technologies allow users to sample music via Twitter, TV, newspapers, websites, and email.[3]

Schedl (2011), for example, explored music listening trends on Last.fm and Twitter, and reported that music popularity across platforms is correlated, so diffusion may involve platform interplay. Few authors have studied the detailed impacts of external news on streaming music diffusion though. To date, Myers et al.'s (2012) work on Twitter is the most complete study of the effects of various types of external news on the diffusion of tweets in Twitter. The authors constructed a diffusion model of tweets over time, and found that 29% of Twitter information propagation is due to external information drivers. They also pointed out the different effects for different types of news information, including sports, business, entertainment, and travel news, among others. The authors only studied the effects in general though, and did not analyse the effects of information content at a finer level of granularity, for example, by assessing music-related news in the entertainment category.

Information that is available from different platforms is known to have different diffusion patterns. Myers et al. (2012) showed that news diffusion on Twitter has rapid information mobility, but music diffusion on a streaming music platform is a longer-term process, since music is a durable information good (Poddar 2006). More useful insights may be related to how long such information affects the choices users make of what music to listen to, and whether there are diverse effects for different types of music information sources. So in this work, we focus on the discovery effects on streaming music diffusion, to assess how external information drives music diffusion in social networks.

[3] As such, it is not easy to determine what is the information source that prompts music diffusion without conducting randomized experiments or user studies. Research in CS has sought to find the correlation between diffusion and external information via observational data and empirical designs, but has not done enough.

Research Setting and Data

In this research, we examined Last.fm, an online music community that integrates music listening, social activities, and social recommendations into a single platform. Besides music streaming, Last.fm also supplies a special "Events" column to broadcast important activity related to a specific artist, such as a coming concert or a live show. Last.fm users can access external information through its platform, and also have separate access to the Internet. To explain streaming music diffusion, we focus on both macro- and micro-level listening changes. For macro-level diffusion, we analyze the effects of external information discovery on an artist's global listening, as well as in specific geographical locations. For micro-level diffusion, we analyze listening changes at the individual user level, including social influence, which represents the effects of social relationships and system recommendations. To gauge the external information discovery effects, we use weekly listening log data to measure social streaming diffusion, summed to the month, when more aggregated observations are necessary.

Data Collection

We acquired listening records for 41 weeks from January 2013 to November 2013 in two stages.

User selection. Via Last.fm, we collected data on 110 most popular artists for 2008 to 2013. They had 100+ million play counts through 2017. For each artist, we also extracted the top fans of the artist's top songs. This yielded 18,933 seed users. They represent those who had already listened at least one of the artists before our observation period, and had the potential to keep listening to the artist in the future.

To obtain a more representative set of users, we also included users' social relationships. Users on Last.fm are linked to each other through their social structure and relationships. Last.fm supplies data for: friends possibly with a real-world social relationship, or with similar listening tastes; and neighbors, who are recommended by Last.fm due to their listening similarities. These users are existing or potential listeners of an artist's music. We extracted 1-hop social relationships for the seed users, and obtained information about 202,966 of their friends and 383,522 of their social network neighbors. Overall, the weekly listening logs of 557,554 users were downloaded for the observation period. The behavior of these users offers a representative snapshot of the basic music listening patterns on Last.fm.[4]

Artist selection. Based on preliminary statistics on the artists to whom users listened, we assessed 10,000+ artists for inclusion. The top 1,300 of them with at least 10 music tracks, including 110 who were popular through 2017, were finally selected as our study targets. We targeted popular artists because it is possible to observe visible diffusion within a short observation period for them.

Panel Dataset Construction

Our raw data contain weekly listening records for 557,554 users. For each user, we captured data on the extent of their listening, *#ArtistListenedTo*, by *Week* and by new *Artist.* On average, each user listened to around 30 different artists weekly; statistics on weekly listening behavior are shown in Figure 1 (left). We also gauged each user's average changes in the different artists they listened to by *Week* and by *Artist* in our 41- week observation period, and calculated the proportion of new artists listened to, Δ:

$$\Delta \# ArtistListenedTo_j = \frac{1}{40} \sum_{t=2}^{41} \frac{\#\ New\ Artists\ User\ j\ Listened\ to\ in\ Wk\ t}{\#\ Artists\ User\ j\ Listened\ to\ in\ Wk\ (t-1)} \tag{1}$$

[4] Based on statistical listening overlaps between seed users and their social relations, there was not very much overlap: only ~9.5% with friends, and ~20.5% with neighbors. This is consistent with Sharma and Cosley's (2016) findings on the over-estimation of social influence on music diffusion.

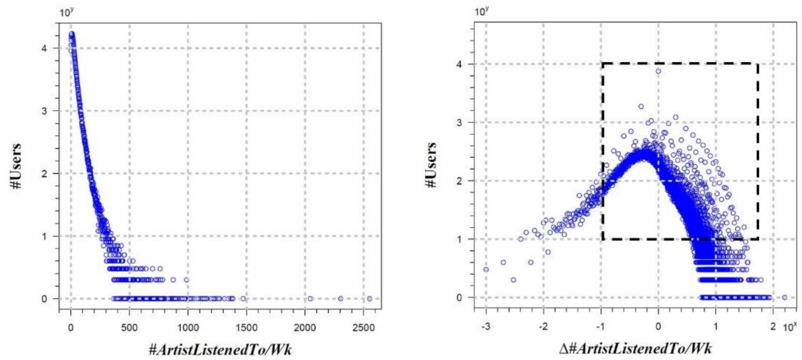

Figure 1. User Listening Behavior, by Artist and by Week

The data plot for Δ*#ArtistListenedTo* indicates that users had large listening weekly change rates (Figure 1, right). Most users listened to at least 10% (Δ*#ArtistListenedTo* > 10^{-1}) new artists compared to the week before (dashed box, Figure 1, right). This rate increased substantially (by 10^x with $x > 0$), which indicates that user listening behavior is dynamic, and the change rate is salient and varied. Dynamic listening suggests the dataset is suitable for diffusion analysis to understand music network users better.

We created two panel datasets at the macro- and micro-levels by merging user listening with artist and user characteristics, social network effects and country information data. Using *propensity score matching* (PSM) (Dehejia and Wahba 2002), we developed balanced panel data subsets for econometric analysis, to estimate music diffusion at the macro-level. Based on a more detailed scan of each user's listening behavior, we also constructed a finer-grained panel dataset for assessing music diffusion at the micro-level. We now show the variables and controls for artists, users and countries (see Table 1).

Table 1. Notation and Definitions of the Study Variables

VARIABLES	DESCRIPTION	VALUE
DEPENDENT VARIABLES		
Artist#Plays/Wk	# times artist's music is played by all users each week	Numeric
Artist#Listeners/Wk	# unique users who listened to an artist's music each week (not Δ change rate)	Numeric
Artist#Plays/Mo	# times an artist's music played by all users each month	Numeric
Artist#Listeners/Mo	# unique users who listened to an artist's music each month (also not Δ)	Numeric
UserArtist#Plays/Mo	# times an artist's music played by a specific user monthly	Numeric
MAIN EFFECTS VARIABLES		
ArtistExtInfoRel	External info release occurred for artist, 1; 0 otherwise	Binary
AfterRelease	Period after artist's external info released, 1; 0 otherwise	Binary
ArtistExtInfoType	Type of external info released on an artist	Category
ExtInfoWeekAfter	Week # (-1, 1, 2, 3, 4) after external info released	Category
CtryExtInfoRel	Country where external info was released (multiple variables)	Binary
ARTIST CONTROL VARIABLES		
LongPopLast.fm	Top chart popularity on Last.fm, from 2005 to 2013	Numeric
LongPopBB	Top chart popularity on Billboard, from 2005 to 2013	Numeric
ShortPopLast.fm	Top chart popularity on Last.fm, 1 month before info release	Numeric
ShortPopBB	Top chart popularity on Billboard, 1 month before info release	Numeric
Artist	Two gender variables, Male (1, 0), female (0, 1) with band (0, 0) as base case	Binary
MajorLabel	Whether artist is connected with major music label	Binary
Genre	Artist's music genre (18-d numeric variable-based genre vector)	Vector
USER CONTROL VARIABLES		
ListeningScale	# of artists user listened to	Numeric
ListeningBreadth	User's diversity of music listening across artists	Numeric
ListeningTaste	User's listening taste (18-d numeric variable-based genre vector)	Vector
TasteSimilarity	Taste similarity for user with artist	Numeric
#Friends	# of friends of user who listened to artist	Numeric
#Neighbors	# of neighbors of user who listened to artist	Numeric
YrsSinceReg	# of years since registration	Numeric
Ctry	Country where user is from	Category
CtryExtInfo	External info released in user's country, 1; 0 otherwise	Binary
Artist#ExtInfoRelease	# of artists with external info in same period, listened to by a user	Numeric

Main Effects Variables: External Information

People access news and event information for artists through various Internet and other selected channels (e.g., Last.fm, Spotify). We capture such changes that affect music diffusion from multiple sources via the Internet. Considering just one kind of external source of information may create bias for geography, culture, information category, etc., we used Google Trends to support the identification of various sources of external information for an artist. Google Trends offers a relative complete sample of search data covering multiple categories, such as Entertainment, News, and other sources like YouTube.

Examining what people search for provides a perspective on their preferences and interests. External sources of information suggest their general interest in a topic, compared to typical search volume. We used weekly change rates for the number of searches, and selected weeks in which rates of change were 50% larger than the prior week. For each external source, we filtered the information by checking its content based on what could be learned from publicly-available data, for example, Wikipedia, Pitchfork, Setlist.fm, Google News, and Last.fm events. We clustered them into two categories: *Music Content Information* and *Non-Music Content Information*, with eight types (see Table 2). *Music Content Information* is directly connected with new music products, including *Single-Song*, *Album*, and *Music-Video Releases*. *Non-Music Content Information* is more diverse, and covers five types of artist social activity: *News, Artist Life* (e.g., birthday, marriage); *News, Music-Related/Music Awards* (e.g., Grammy Awards, news of a coming album); *Tour/Concert*; attending *Live TV Shows* (e.g., Saturday Night Live); and *Live Performances* at music festivals. Such external information may not be directly connected with new music products, but still may attract people's attention when they are reading the news or watching TV.

Some artists had one instance of external information released in our study period, while others had multiple releases: for example, they may arrange a *Single-Song Release* (Type 6), then an *Album Release* (Type 7), followed by a *Concert* (Type 3) week by week. To reduce the effects of multiple external information at the same time, we used only one external information for each artist during a week. For artists with multiple releases, we selected the one that had at least a two-month gap from others that were identified, to reduce possible over-estimation of the effect of a release. For the 1,300 artists, 407 had external source-based information releases during the study period, 210 had new *Music Content Information*, and 197 released new *Non-Music Content Information* (see Table 2).

Table 2. External Information Source Type (*ArtistExtInfoType*)

TYPE	DESCRIPTION	# ARTISTS
Non-Music Content Information		**197**
1	*News, Artist Life*	48
2	*News, Music-Related Info, Music Awards*	47
3	*Tour, Concert*	40
4	*Live TV Show*	28
5	*Live Performance / Festival*	34
Music Content Information		**210**
6	*Single-Song Release*	66
7	*Album Release*	131
8	*Music-Video Release*	13

Time and geographical variables. Streaming music listening should not be bound by the time or any limits of geography, due to the ubiquitous nature of the Internet. However, the effect of an instance of external news may affect music listening and be limited both by the timing of its release and the country where it was released, especially for *Non-Music Content Information*, such as external information on TV shows and concerts. A strong regional event may only have impact on local listeners, for example.

Understanding the type, time effect and geolocation of external information may offer deeper insights into the music labels for what kinds of news and information can be used for music promotion.

Dependent Variables: Streaming Music Diffusion at the Macro- and Micro-Levels

Macro-level. Music social networks often use the number of times an artist's music is played (*Artist#Plays*) and the number of users who listen to a unique artist's music at least once (*Artist#Listeners*) to rank their popularity. *Artist#Plays* reflects an artist's general popularity, while the latter indicates the artist's market penetration with users. YouTube uses *Artist#Plays* to measure track popularity; and Last.fm and Spotify track both *Artist#Plays* (scrobbles) and *Artist#Listeners*. Some users listen to a song once, while others listen more, so these dependent variables are complementary for assessing diffusion. We use both to measure each week's music diffusion at the macro-level, and the monthly diffusion for a robust check. Statistics of diffusion are shown in Table 3. We note that *Artist#Plays* is five times as large as *Artist#Listeners*. On average, users listened to an artist five times.

Table 3. Streaming Music Diffusion Statistics with Averages for January to November 2013

STREAMING MUSIC DIFFUSION	MIN	MAX	MEAN	SE
WEEKLY				
Artist#Plays/Wk	0.05	1,483.3	14.5	44.3
Artist#Listeners/Wk	0.03	85.7	2.8	5.0
MONTHLY				
Artist#Plays/Mo	0.45	4,649.8	59.2	172.3
Artist#Listeners/Mo	0.14	309.2	11.2	20.3
Notes. 53,300 obs., 41 obs. per artist for 1,300 artists in 000s of tracks streamed.				

Micro-level. Music diffusion at the micro-level is measured with *UserArtist#Plays*. This represents how many times a specific user listened to an artist after news information on the artist was released through an external source (see Figure 2). To eliminate the possible endogeneity effects, the users we selected for analysis could not have had listening records for a given artist prior to the release of external information, but may have had listening records thereafter. We observed users' listening behavior for two months before the *ArtistExtInfoRelease* and one month afterwards, for the two different periods. We focused on the one-month period after an external release of information occurred. This is because it is not possible to guarantee that there are no influence effects mixed into the discovery effects that users experience, if the observation time is too long. We mitigated this kind of outcome by observing the effects of external information release in a limited 4-week time period after it occurred.

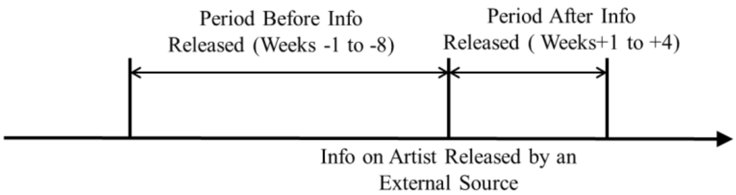

Figure 2. Observation Periods for Music Diffusion at the Micro-Level

Control Variables in Propensity Score Matching and Modeling

Music diffusion is affected by multiple factors, including artist charateristics, user diversity, the competitive environment of the music market, and so on. To obtain a balanced panel dataset that can be used in empirical testing for the effects of external information on music diffusion, we considered various control variables, and applied the PSM procedure to match observations on the basis of statistical matches. This accounts for artist and user heterogeneity in music characteristics and listening behavior.

Artist. Artist music listening is influenced by a number of factors:

- **Popularity.** An artist's popularity contributes to the diffusion of their music (Ren and Kauffman 2017). We refer to this as *artist popularity*, and it is possible to distinguish the length of time that it lasts. Popularity exhibits an accumulating process and is dynamic over time, so we use long-term and short-term popularity to describe it. *Long-term popularity* (*LongPop*) is the accumulated level of popularity from 2005 until 2013, just before our study period. *Short-term popularity* (*ShortPop*) is defined as its observed level during the month prior to the time the observation week occurred. Both measures are based on the number of appearance on the Billboard Hot 100 and Last.fm weekly top-100 ranking charts at the relevant times.
- **Artist.** Two binary variables indicate male (1, 0) or female (0, 1), with band (0, 0) as base case.

- **Major label.** One binary variable represents whether an artist had a major music label (1) or not (0). Three *MajorLabels* (Universal, Sony/EMI, Warner) operated during our study period.
- **Genre.** A *support vector machine* (SVM) was used to learn an 18-dimensional music genre vector from the acoustic content of the artist's top songs and Last.fm tags, for genre probability.

User. A user's choice of what to listen to in a social network may be affected by multiple factors:

- **User listening behavior**. This includes how much and how diverse a user's music listening behavior was before information release occurred. *ListeningScale* is the number of unique artists a user listened to, and *ListeningBreadth* indicates how much a user listened to different artists (Garg et al. 2011). We used the *Gini coefficient* for the artist-level distribution of the number of songs to measure listening diversity. We also used *Euclidean distance* for the *TasteSimilarity* for a user and an artist with the user's current taste and the artist's genre.
- **Social influence.** This is represented by a continuous variable for the number of a user's social friends or neighbors who listened to the artist, and may have affected their listening choices.
- **External information released.** We included two variables to indicate that news about an artist represents different types of information. One is *ArtistExtInfoType*, representing the base case (Type 1) and seven other types. In addition, for all the artists with external information releases in the same observation period, we measured the number of unique artists whose music each user listened to, *Artist#ExtInfoRelease*. This measure suggests a possible *crowding effect* that makes any individual external information release have less impact.
- **Geolocation.** The binary variable *CtryExtInfo* codes for whether an artist's information release or an artist's event (concert, TV show) was in the country of the music-listener.

Model and Methodology

We applied a *difference-in-differences* (DiD) model (Imbens and Wooldridge 2007) to examine the causal effects of external information on music diffusion. Our model involves a hierarchical empirical method with steps from the general to a more fine-grained perspective. Prior to estimation, we applied PSM to address artist and user endogeneity for music listening choices and music diffusion. We constructed a macro-level control group, in which artists had no external information released in a common observation period, but where there were comparable artist characteristics and historical music diffusion in the treatment group. A fine-grained control group was also constructed by consider the user's listening behavior and a geolocation effect. By studying the individual user-level, we hope to learn about the effect of external information on how the content of a user's music listening activities is changed.

PSM to Address Artist and User Endogeneity in Music Diffusion

Artist. For each treated artist with external information released in our study period, we applied a time-dependent propensity score match with a control artist who had no external information released during the same week. For this, we estimated a logit model at the artist-level to obtain a balanced distribution of the observed covariates for the treatment and control group observations, including artists' characteristics, such as music genre, and long-term and short-term popularity.

We also balanced the artists' previous listening status one month before external information was released. This allowed us to observe the direct feedback on music diffusion for two similar artists, with and without external information in the same observation period.

$$\begin{aligned}\Pr(ArtistExtInfoRel_{it} = 1 \mid \cdot) = f(&Artist_i,\ MajorLabel_i,\ Genre_i, LongPopLast.fm_i,\\ &LongPopBB_i,\ ShortPopLast.fm_{it},\ ShortPopBB_{it},\\ &Artist\#Plays/Mo_{i,t-1},\ Artist\#Listeners/Mo_{i,t-1}\ , t) \qquad (2)\end{aligned}$$

Finally, we selected 407 control artist observations from the total of 35,999 control observation for artists. The PSM results are summarized in Table 4, which shows that the treatment and control artists are properly matched. These 814 artists were further used to analyze music diffusion at the macro-level.

Table 4. Propensity Score Matching Results for Music Artists

ARTIST CONTROL VARIABLES	TREATMENT	CONTROL CANDIDATE	CONTROL MATCHED
Artist#Plays/Mo	137,181	99,778	149,486
Artist#Listeners/Mo	23,209	16,689	23,290
Artist: Male	0.248	0.199	0.248
Artist: Female	0.155	0.102	0.194
MajorLabel	0.482	0.440	0.514
Music Genres			
Rock	0.618	0.684	0.601
Alternative	0.222	0.256	0.225
Indie	0.284	0.346	0.259
Pop	0.269	0.204	0.260
Hip-hop	0.067	0.073	0.059
Rap	0.033	0.033	0.027
R&B	0.048	0.020	0.065
Electronic	0.119	0.115	0.117
Metal	0.220	0.220	0.217
Folk	0.081	0.076	0.102
Soul	0.041	0.036	0.042
Experimental	0.087	0.001	0.089
Punk	0.041	0.093	0.045
Classic	0.014	0.047	0.006
Jazz	0.025	0.019	0.029
Blues	0.034	0.021	0.036
Country	0.004	0.014	0.010
Reggae	0.011	0.007	0.048
LongPopLast.fm	25.830	33.680	22.510
LongPopBB	20.290	11.430	23.640
ShortPopLast.fm	0.378	0.196	0.415
ShortPopBB	0.315	0.086	0.354

Note. Numeric entries represent the statistic mean of each control variable, used for artist matching. The *Control Candidate* column is the statistical result of 35,999 observation for artists who had no external information releases in the study period. *Treatment* and *Control Matched* are the statistics for the 407 artists in each group.

User matches based on geographic location. Some of the external information has obvious geographical bounds for its informational relevance, such as TV Shows, Music Festivals, and Concerts, which occur in a specific country or city. To learn whether there are geographic restrictions on music diffusion related to external information releases, we further constructed a fine-grained panel dataset according to the country where the external information was released. Among the 407 artists in our treatment group with information released, 199 had external information releases in the U.S. and 76 in the U.K. We focused on these two top-ranked countries for comparison. There are 46,200 users in our dataset that are in the U.S., and 18,402 from the U.K., out of the total users. We tracked the listening change differences for U.S. and non-U.S. users, and for U.K. and non-U.K. users. Similar to our artist treatment-and-control groups, we used the users' country, *Ctry*, to do the matching, by balancing their registration time and listening behavior, as follows:

$$\Pr\left(Ctry_j = 1 \middle| \cdot\right) = f(RegSinceYear_j,\ ListeningTaste_j,\ ListeningScale_j,\ ListeningBreadth_j) \quad (3)$$

This resulted in 92,400 users tracked for U.S. external information releases, and 36,804 for U.K.

Difference-in-Differences (DiD) Model at the Macro-Level

We examined the effects of external information on music diffusion at the macro-level with a DiD model. This is ideal for the structure of our analysis. Music diffusion at the artist level *i* is represented with a pre-and-post-DiD model:

$$\begin{aligned} DepVar_{it} = Constant &+ \beta_1 ArtistExtInfoRel_{it} + \beta_2 AfterRelease_{it} \\ &+ \beta_3 ArtistExtInfoRel_{it} \times AfterRelease_{it} \\ &+ \beta_4 ArtistExtInfoType_{it} + \beta_5 ExtInfoWeekAfter_{it} + \epsilon_{it} \end{aligned} \quad (4)$$

The $DepVar_{it}$ is for music diffusion, represented by weekly or monthly *Artist#Plays* and *Artist#Listeners* for artist *i* at observation week *t*. *ArtistExtInfoType* is the type of the external information, we use

Type 1 (*News, Artist Life*) as the base case, because of it has a distant relationship with the music product itself, and can be compared with other external information. This main effect variable indicates the validity of the effects of various types of information. *ExtInfoWeekAfter* indicates the week after external information was released related to an artist. The listening records for users in the week prior to the external information release represent the basis for comparison. We can see the change in music diffusion after external information was released compared to before that. We next show the music diffusion changes at the macro-level for all users, and also at a more fine-grained level by considering user geolocation information separately.

Count Data Model for Micro-Level Analysis

At the micro-level, we modeled the listening change of a user after an artist's external information is released. We wish to know how much external information will affect an individual's listening behavior. As Figure 2 shows, we tracked users with no record of listening to an artist before the artist's information was released, but may have listened to them after that. We examined the effects of user listening taste, and social network and external effects. So we considered seed users with observable social relations. The music diffusion of artist *i* to user *j* at time *t* is represented with a count data model:

$$\begin{aligned} UserArtist\#Plays_{ijt} = {} & Constant + \beta_1 ArtistExtInfoType_{it} + \beta_2 Artist_i + \beta_3 MajorLable_i + \beta_4 Genre_i \\ & + \beta_5 LongPopLast.fm_i + \beta_6 LongPopBB_i + \beta_7 ShortPopLast.fm_{it} + \beta_8 ShortPopBB_{it} \\ & + \beta_9 ListeningBreadth_{jt} + \beta_{10} ListeningScale_{jt} + \beta_{11} TasteSimilarity_{ijt} + \beta_{12}\#Friends_{ijt} \\ & + \beta_{13}\#Neighbors_{ijt} + \beta_{14} Artist\#ExtInfoRelease_{jt} + \beta_{15} CtryExtInfo_{ijt} + \epsilon_{ijt} \end{aligned} \quad (5)$$

For each music diffusion model, we used a *negative binomial* count data model. It is suitable for our study since the dependent variable has non-negative values of 0, 1, etc.. Also, it is suitable when *over-dispersion* is present, that is the variance of the mean count is higher than the observed theorized count.

Results and Interpretation

We present the estimation results for the DiD and count data models next. The results include external information effects on music diffusion at the macro- and geographic levels, scale and persistence of various types of information effects, and micro-level listening changes.

Music Diffusion at the Macro-Level

We estimated the listening changes of 557,554 users for the 814 matched artists using Eq. 4, to test whether an external information release had a positive impact on monthly and weekly music diffusion on total *Artist#Plays* and *Artist#Listeners.* Table 5 gives the DiD regression results.

Table 5. DiD Regression Results for Music Diffusion at the Macro-Level

Main Effect Variables	*Artist#Plays / Mo* (I) (SE)	*Artist#Listeners / Mo* (II) (SE)	*Artist#Plays / Wk* (III) (SE)	*Artist#Listeners / Wk* (IV) (SE)
Constant	11.39 *** (0.06)	9.59 *** (0.05)	9.87 *** (0.05)	8.00 *** (0.05)
ArtistExtInfoRel	0.05 (0.06)	0.03 (0.05)	-0.21 (0.04)	-0.05 (0.03)
AfterRelease	-0.03 (0.06)	-0.03 (0.05)	-0.02 (0.03)	-0.01 (0.02)
ArtistExtInfoRel × *AfterRelease*	0.40 *** (0.08)	0.14 * (0.07)	0.34 *** (0.04)	0.13 *** (0.03)

Notes. Model: Neg. bin.; mo.=month, we.=week; 1,628 mo. obs. for I, II = (407 + 407) × 2; 6,512 wk. obs. for III, IV = (407 + 407) × 8. Pseudo-R^2: I – 37.6%, II – 36.4%, III – 44.9%, IV – 45.6%; shape parameter, α: I – .66, II – .52, III – .60, IV – a.46. Signif.: * $p < .10$; ** $p < .05$; *** $p < .01$.

In all models, *AfterRelease* was negative but not significant. A possible reason is that the control group had a larger listener base than the treatment group, as shown in Table 4. So the increase was not enough to produce an average change for all 814 artists. This coefficient might have been positive if the treatment group's listener base were equal to or greater than the control group's. The coefficient of *ArtistExtInfoRel* was positive for the monthly dependent variables, but negative for the weekly ones. This occured for the comparison between the treatment and control groups, before and after external information was released. So on a monthly basis, the treatment group experienced more music diffusion than the control group. When we made a weekly comparison, the treatment group had no greater evidence of music diffusion. This indicates that the effect of external information may be limited as time passes.

Regardless of which dependent variable we used, the targeted main effect *ArtistExtInfoRel* × *AfterRelease* was positive ($p < 0.01$). This suggests that external information release was not associated with a decline in music diffusion across the treatment and control groups, before and after the external information release. If we examined monthly music diffusion, compared to the control group, the treatment group had a 49.2% increase in *Artist#Plays*, and a 15.0% increase in *Artist#Listeners* when an external information release occurred (*Artist#Plays*: ($e^{0.40}$ – 1) = 49.2%, *Artist#Listeners*: ($e^{0.14}$ - 1) = 15.0%). Weekly music diffusion had a similar increase.

Although the DiD regression results in Table 5 indicate a significant impact of external information, the result was the average for all types of external information releases. However, the extent to which *Non-Music Content Information* affects diffusion is not clear. So we further tested the effect of *Music Content* and *Non-Music Content Information* separately according to Eq. 4, but only focused on the 407 artists who had external information. Table 6 shows the results for weekly music diffusion.

The results show the effects and persistence over time of *Music Content* and *Non-Music Content Information.* Type 1, *News*, *Artist Life*, was the base case for external information release. Not surprisingly, all types of *Music Content Information* led to a significant increase in the *Artist#Plays* and *Artist#Listeners* dependent variable values. Among the three types, Type 8, *Music-Video Release*, resulted in the largest change in diffusion. This indicates that people are more attracted by 3D videos or stories compared to voice only, and also suggests why the music industry invests a lot in MTV videos. In addition, *Music Content Information* had a persistent effect on music diffusion in the month after information was released, and although it lessened over time, the trend was still increasing.

Non-Music Content had some dissimilar effects compared to *Music Content Information*. Not all types had significant increases, though some led to music diffusion increases for both dependent variables. For example, Type 4, *Live TV Show*, had the highest effect, with 36.3% in *Artist#Plays* and 41.9% in *Artist#Listeners*. Type 3, *Tour*, *Concert*, seemed to have increased total playing time, but did not result in music diffusion to new listeners. A possible reason is that a tour and concerts are more likely to attract existing listeners, not new ones. Another interesting finding is that the persistence of the effect over time of *Non-Music Content Information* was only about 2 weeks after external information release occurred, although the coefficients for all 4 weeks after the event were positive. Compared to *Music Content Information*, the effect was smaller (0.12 < 0.47, and 0.05 < 0.15, respectively). This further verified why *ArtistExtInfoRel* for *Non-Music Content Information* was negative for weekly diffusion, when we considered all 4 weeks after the external information released (see Table 5).

Table 6. Negative Binomial Regression Count Data Model Results for External Information

MAIN EFFECTS VARIABLES	MUSIC CONTENT INFO *Artist#Plays / Wk* (I) (SE)	MUSIC CONTENT INFO *Artist#Listeners / Wk* (II) (SE)	NON-MUSIC CONTENT INFO *Artist#Plays / Wk* (III) (SE)	NON-MUSIC CONTENT INFO *Artist#Listeners / Wk* (IV) (SE)
Constant	9.23 *** (0.12)	7.35 *** (0.10)	9.63 *** (0.16)	7.94 *** (0.14)
ArtistExtInfoType				
News-Artist Life	Base case	Base case	Base case	Base case
News-Music-Related Info			0.10 * (0.13)	0.14 *** (0.05)
Tour, Concert			0.28 *** (0.06)	0.06 (0.05)
Live TV Show			0.31 *** (0.06)	0.35 *** (0.06)
Live Performance / Festival			0.09 (0.06)	-0.06 (0.06)
Single-Song Release	0.38 *** (0.07)	0.24 *** (0.06)		
Album Release	0.69 *** (0.06)	0.23 *** (0.05)		
Music-Video Release	0.96 *** (0.11)	0.84 *** (0.09)		
ExtInfoWeekAfter				
WeekAfter-1	Base case	Base case	Base case	Base case
WeekAfter1	0.47 *** (0.07)	0.15 *** (0.06)	0.12 * (0.06)	0.05 * (0.06)
WeekAfter2	0.51 *** (0.07)	0.19 *** (0.06)	0.11 * (0.06)	0.04 (0.06)
WeekAfter3	0.35 *** (0.07)	0.11 *** (0.06)	0.03 (0.06)	0.02 (0.06)
WeekAfter4	0.23 *** (0.07)	0.08 * (0.06)	0.001 (0.07)	0.01 (0.06)

Notes. Model: Neg. bin.; total obs. = 2,035; 985 *Non-Music Content Info* wk. obs. = 197 × 5; 1,050 *Music Content Info* wk. obs. = 210 × 5. Type 1, *News, Artist Life*: base case is *ArtistExtInfoType*. *WeekAfter*-1: base case is *ExtInfoWeekAfter*. We compare music diffusion for 1 wk. before and 4 wks. after info released. Shape parameters α: I – .57, II – .41, III – .42, IV – .11; pseudo-R^2: 1 – 48.7%, II – 49.6%, III – 55.6%, IV – 55.9%. Signif: * $p < .10$; ** $p < .05$; *** $p < .01$.

Diffusion Diversity at the Geographic Level

Some external information releases have obvious geographical bounds for their relevance, especially *Non-Music Content* Information. We tested for diffusion diversity at the geographic-level with the fine-grained panel dataset, via this additional model, based on Eq. 4.

$$\begin{aligned} DepVar_{it} = Constant &+ \beta_1 CtryExtInfoRel_{it} + \beta_2 AfterRelease_{it} \\ &+ \beta_3 CtryExtInfoRel_{it} \times AfterRelease_{it} + \beta_4 ArtistExtInfoType_{it} \\ &+ \beta_5 ExtInfoWeekAfter_{it} + \beta_6 Artist_i + \beta_7 MajorLable_i \\ &+ \beta_8 Genre_i + \beta_9 LongPopLast.fm_i + \beta_{10} LongPopBB_i + \epsilon_{it} \end{aligned} \tag{6}$$

Our treatment group includes 46,200 U.S. users, and the control group has a matched set of 46,200 non-U.S. users (we did the same analysis for the U.K. users, but omitted here because of the space limitations). We changed the treatment group variable from *ArtistExtInfoRel* to *CtryExtInfoRel*. The dependent variables represent the cumulative listening counts for the treatment and control groups related to 199 artists with external information released in the U.S. The results are shown in Table 7.

Table 7. DiD Regression Results for External Information at the Geographic-Level

MAIN EFFECT VARIABLES	MONTHLY		WEEKLY	
	MUSIC CONTENT (I) (SE)	NON-MUSIC CONTENT (II) (SE)	MUSICCONTENT (III) (SE)	NON-MUSIC CONTENT (IV) (SE)
Constant	7.85 *** (0.27)	7.79 *** (0.27)	4.87 *** (0.15)	5.11 *** (0.14)
CtryExtInfoRel	0.26 *** (0.08)	0.16 * (0.08)	0.29 *** (0.07)	0.21 *** (0.07)
AfterRelease	0.48 *** (0.08)	0.04 (0.08)	0.10 * (0.06)	0.02 (0.06)
CtryExtInfoRel × *AfterRelease*	0.06 (0.11)	0.08 (0.39)	0.03 (0.08)	0.03 (0.08)
ArtistExtInfoType				
News-Artist Life	Base case	Base case	Base case	Base case
News-Music-Related Info		0.63 *** (0.15)		0.27 *** (0.08)
Tour, Concert		-0.01 (0.15)		-0.05 (0.08)
Live TV Show		0.45 *** (0.15)		0.43 *** (0.07)
Live Performance / Festival		0.08 (0.14)		-0.09 (0.07)
Single Song Release	0.24 * (0.13)		0.26 *** (0.07)	
Album Release	0.49 *** (0.12)		0.31 *** (0.06)	
Music Video Release	0.96 *** (0.19)		0.84 *** (0.10)	
ExtInfoWeekAfter				
WeekAfter-1			Base case	Base case
WeekAfter1			0.15 ** (0.05)	0.12 * (0.05)
WeekAfter2			0.28 * (0.05)	0.11 (0.05)
WeekAfter3			0.11 * (0.05)	0.04 (0.05)
WeekAfter4			0.005 (0.06)	0.001 (0.06)

Notes. Overall model obs.: I – 512, II – 376, III – 1,280, IV – 940. 105 artists had *Music Content Info*; 94 had *Non-Music Content Info*; 512 mo. obs. = (105 + 23 base case) × 4; 1,280 wk obs. = (105 + 23 base case) × 10. Shape parameters α: I – .42, II – .46, III – .29, IV – .21; pseudo-R^2: 1 – 68.4%, II – 62.6%, III – 66.5%, IV – 74.5%. I only show regression results for U.S. *Artist#Plays* for treatment and control groups. *Artist#Listeners* had similar results, as did the U.K. data. I omitted the U.K. and other control variable results due to space limitations. Signif: * $p < .10$; ** $p < .05$; *** $p < .01$.

Across the models that were used, the treatment variable *CtryExtInfoRel* was positive and significant. This confirms the implied hypothesis, that music diffusion has geographic bounds, not just for *Non-Music Content Information,* but also for *Music Content Information.* The coefficient estimates for *AfterRelease* were all positive, but only significant in the case of *Music Content Information.* The reason is that *Music Content Information* seems to have had a positive impact on music diffusion, but the *Non-Music Content Information* variables were not as consistent in their estimated effects. The coefficients for *CtryExtInfoRel* × *AfterRelease* also were positive but not significant ($p = 0.43$ for the monthly *Music Content Information* variables; and $p = 0.48$ for the weekly *Music Content Information* variables – both in the DiD regression). A possible reason is that the existing user listening diversity across the selected 199 artists with external information was relatively high (*Artist#Plays/Mo,* mean = 17,730, SE = 31,008; for *Artist#Plays/Wk*, the mean is 9,216, the SE is 14,117). So the average increase in diffusion was not big to be significant. However, the evidence for positive additional diffusion is a further verification of the geographic effect. The effects of *ArtistExtInfoType* and *ExtInfoWeekAfter* had similar results as the music diffusion which occurred at the macro-level.

Listening Diversity at the Micro-Level of Users

We next analyze music diffusion at the user-level. We focused on new listeners of the 407 artists 1 month after external information was released. Table 8 shows our results for listening diversity. *CtryExtInfo* was positive and significant, so music diffusion was geographically bounded. For the social effects, even when we considered all of the users' social relations for those who adopted an artist's music, the effect on user listening choices was still smaller than the user's own listening behavior (*#Neighbors* = 0.04 < *ListeningBreadth* = 0.70 < *TasteSimilarity* = 1.10, p <0.01). External information releases had similar effects to what we observed at the macro-level. *Artist#ExtInfoRelease* assessed what happened when many artists released information at the same time, and this negatively impacted an artist's music diffusion (-0.09; $p < 0.01$). Thus, the music industry must select a suitable time to release albums, to mitigate competitive effects from other artists.

Overall, the results indicate that external information discovery, user geolocation and listening, and social influence should be considered for the design of effective, personalized music recommendations.

Table 8. Count Data Regression Results for the Micro-Level Analysis

VARIABLES		VARIABLES	
Constant	1.38 *** (0.05)	***ArtistExtInfoType***	
Geographic Effect		*News-Artist Life*	Base case
CtryExtInfo	0.38 *** (0.04)	*News-Music-Related Info*	0.21 *** (0.03)
User's Listening Behavior		*Tour, Concert*	-0.04 (0.04)
ListeningScale	0.004***(0.00)	*Live TV Show*	0.01 * (0.03)
ListeningBreadth	0.70 *** (0.06)	*Live Performance / Festival*	-0.05 (0.03)
TasteSimilarity	1.10 *** (0.07)	*Single Song Release*	0.21 *** (0.03)
Artist Characteristics		*Album Release*	0.48 *** (0.03)
MajorLabel	-0.10 ** (0.02)	*Music Video Release*	0.03 * (0.05)
LongPopLast.fm	0.003** (0.00)	**Social and Crowding Effects**	
LongPopBB	0.004 ** (0.00)	*#Friends*	0.003*** (0.00)
ShortPopLast.fm	0.07 ** (0.02)	*#Neighbors*	0.04 *** (0.00)
ShortPopBB	0.08 ** (0.02)	*Artist#ExtInfoRelease*	-0.09 *** (0.00)

Note. Model: Neg. bin., 62,000 user-level listening obs. on 407 artists who had external info. Pseudo-R^2 = 24.9%; shape parameter, $\alpha = 1.80$. Signif: * $p < .10$; ** $p < .05$; *** $p < .01$.

Discussion and Conclusion

Music labels release music for free listening before CDs are released, to attract attention from existing and new consumers. Semi-closed music social networks encourage sharing of social information, but are open to external information discovery by their users. In this complex environment, it is important for music social network providers to understand how the diffusion of music works in their business models. They may find that there is a hidden source for consumer engagement and higher profit.

We investigated how diffusion of an artist's music is affected when new external information is released outside of a music social network. By analyzing Last.fm data, we can offer managerial insights that ought to be useful for music promotion and personalized recommendations in online music platforms. First, the discovery of external information from multiple channels by users had a positive impact on reversing the decline in streaming music listening, and different kinds of information exhibited different impacts. New *Music Content Information* appeared to make it easier to attract new listeners in a relatively short time. *Non-Music Content Information* is less effective in comparison to *Music*. Mass media including TV and newspapers can also encourage people to actively search and listen to new artists.

Second, how to attract new listeners and keep the current listeners' attention between new music releases is an important issue for music labels in the industry. The related time period is key for *Non-Music Content Information*, because it helps retain users when music promoters recommend an artist.

Third, streaming music diffusion still has geolocation bounds though, so external information is more likely to attract the attention of local listeners. Although people can access whatever music they like online, their more limited access to external information may limit their choices. This is true for both *Music* and *Non-Music Content Information*. So location targeting is key for information discovery.

Fourth, for recommendations, external information, user geolocation, and listening behavior can be leveraged to improve personalization. For consumers, our study suggests to improve personalized music recommendations, by leveraging information from various channels for the platform. For the industry, we also offer managerial insights on target consumer selection for more effective artist promotion.

There are some limitations though. It is hard to capture all relevant information for artists. We only considered a single channel for an artist. Some had frequent releases that we could not observe. Others arranged tours after a song release, for more local diffusion. The effect of a new song should reflect the cumulative impact of external information releases. This probably did not affect our estimation results for *Non-Music Content Information*, beyond slight over-estimation of the impact of the effect of *Music Content Information*. Also, we cannot guarantee no mixing of influence and discovery: we only observed the diffusion of music over a limited time. To distinguish these effects, we need data over a longer period, and to consider effects of social relations and capital, and weak and strong social ties.

Our estimation work was performed on a subset of Last.fm's data, so there may be selection bias. We plan to study music diffusion at the macro-level across the platform, to comprehensively assess the effects of external information discovery. We only considered whether users were from the country where an artist's external information was released. For control group users, we did not do more finer clustering (e.g., country traits, including language differences, cultural and physical distance). U.S. artists may find it is easier to attract U.K. listeners than China's, due to language and cultural similarities. We are currently assessing the addition of these distance variables to more deeply analyze the geographic effects.

References

Bakshy, E., Rosenn, I., Marlow, C., Adamic, L., 2012. "The Role of Social Networks in Information Diffusion," in Proc. 21st Intl. Conf. on World Wide Web, New York: ACM, pp. 519-528.

Bapna, R., Umyarov, A. 2015. "Do Your Online Friends Make You Pay? A Randomized Field Experiment on Peer Influence in Online Social Networks," *Mgmt. Sci.* (61:8), pp. 1902-1920.

Dehejia, R.H., and Wahba, S. 2002. "Propensity Score-Matching Methods for Non-Experimental Causal Studies," *Rev. Econ. Stat.* (84:1), pp. 151-161.

Dewan, S., Ho, Y., Ramaprasad, J. 2017. "Popularity or Proximity: Characterizing the Nature of Social Influence in an Online Music Community," *Info. Sys. Res.* (28:1), pp. 117–136.

Ellison, N.B., Steinfield, C., Lampe, C., 2011. "Connection Strategies: Social Capital Implications of Facebook-Enabled Communication Practices," *New Media & Soc.* (13:6), pp. 873-892.

Garg, R., Smith, M.D., Telang, R. 2011. "Measuring Information Diffusion in an Online Community," *J. Mgmt. Info. Sys.* (28:2), pp. 11-38.

IFPI . 2012, 2015, 2017. "Digital Music Report: Recording Industry in Numbers," London, UK.

Imbens, G., Wooldridge, J. 2007. "What's New in Econometrics? Difference-in-Differences Estimation," Lecture Notes, NBER Summer Institute, Cambridge, MA.

Myers, A., Zhu, C., Leskovec, J. 2012. "Information Diffusion and External Influence in Networks," in *Proc. 18th Intl. Conf. Knowl. Disc. and Data Mining*, New York: ACM, pp. 33-41.

Pálovics, R., Benczúr, A. 2015. "Temporal Influence over the Last.fm Social Network," *Soc. Netw. Anal. Mining* (5:1), pp. 4-11.

Poddar, S., 2006. "Music Product as a Durable Good and Online Piracy," *Rev. Econ. Res. Copyright Iss.* (3:2), pp. 53-66.

RIAA (Recording Industry Assoc. of Amer.). 2016, 2017. https://www.riaa.com/u-s-sales-database/.

Ren, J., Kauffman, R.J., 2017. "Understanding Music Track Popularity in a Social Network." in *Proc. 2017 Euro. Conf. on Info. Sys.*, Atlanta: Assoc. for Info. Sys., pp. 374-388.

Schedl, M. 2011. "Analyzing the Potential of Microblogs for Spatio-Temporal Popularity Estimation of Music Artists," in *Proc. Intl. Joint Conf. on Artif. Intell.*, Menlo Park: AAAI Press, pp. 539-553.

Sharma, A., Cosley, D. 2016. "Distinguishing between Personal Preferences and Social Influence in Online Activity Feeds," in *Proc.Conf. Comp.-Supp. Coop. Work*, New York: ACM, pp. 1091–1103.

Susarla, A., Oh, J.H., Tan, Y. 2012. "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube," *Info. Sys. Res.* 23(1), pp. 23-41.